PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	How do people perceive different labels for rotator cuff disease? A content analysis of data collected in a randomised controlled experiment
AUTHORS	Zadro, Joshua; Michaleff, Zoe; O'Keeffe, Mary; Ferreira, Giovanni; Haas, Romi; Harris, Ian; Buchbinder, Rachelle; Maher, Christopher

VERSION 1 – REVIEW

REVIEWER	S Carlfjord Linköping University, Dept of Medical and Health Sciences
REVIEW RETURNED	03-Jun-2021

GENERAL COMMENTS	This is an interesting study, using innovative methods for data collection, it is well written and provides valuable insights in how to handle shoulder pain patients. Before being published, there are however some minor corrections that I believe should be made. Below you find a detailed list of comments on your manuscript. Strengths and limitations of the study: Line 51. You state that you used “high-quality methods”, which may be true, but in my opinion these are “standard procedures”. If you thoroughly followed these standardized procedures it is a strength, but you should not claim that they are high-quality methods. Introduction Lines 97-101 in your Introduction is more of a discussion. I suggest that you remove this part. Materials and methods Participants and recruitment Line 112. You mention Qualtrics, with no further explanation here (you do explain it later in text). I suggest that you provide this information to the reader here, when first mentioning Qualtrics. Data collection Line 129. Participants were randomised to the six groups. You earlier described three groups of participants based on experience of shoulder pain – was this randomisation performed within each of these groups? From your results I guess that it was, but I suggest that you make this clear also in the Methods section. In line with this, figure 1 indicates that randomisation was not based on the three groups, which I find confusing. Please, clarify. Data analysis Line 159. I have difficulties with your reference 24. I expected literature on Content analysis, for example Content analysis: an introduction to its methodology by Klaus Krippendorff. I suggest that you add this or a similar reference to the method applied. Line 160. I suggest that you provide some information about the coders here, that they both are experienced physiotherapists… (appears later, but inform the reader at this stage).

	Results I find the results clear and presented in a sufficient way. Discussion Summary of key findings Line 238. Your summary of key findings is relevant, but I find I too long, too much repetition of results. Try to shorten it, and do not repeat numbers and proportions that can be found in the Results section. Strengths and weaknesses Line 265. You present some weaknesses, but did you consider the fact that your population is “selected”. Only those registering in Qualtrics are eligible. Who are you missing? Older people? Poor people? I suggest that you reflect also on this limitation. Comparison to literature Lines 322-327. This is an interesting discussion, but you do not need all these numbers and proportions to understand the message. All these data can be found in the Results section, and the readability increases if you remove them from here. Conclusion I agree with your conclusion, but I find it too long, and again repeating the results. Try to make it more distinct.

REVIEWER	Carlos Torrens Hospital del Mar
REVIEW RETURNED	12-Jun-2021

GENERAL COMMENTS	Cristina Barrufet has helped completing the review. Cristina Barrufet Health Services Evaluation and Clinical Epidemiology Department, Institut Hospital del Mar d’Investigacions Mèdiques (IMIM), Hospital del Mar, 08003 Barcelona, Spain Health Services Research Unit, IMIM-Institut Hospital del Mar d'Investigacions Mèdiques, Barcelona, Spain Study design Has been used any qualitative methodology approach (phenomenology, hermeneutics, grounded theory)? It should be necessary to define the level of interpretation (descriptive, descriptive-interpretative, etc) and the perspective / level of participation (ETIC). Has been considered sufficiency and pertinence of the sample based on the analysed qualitative data? Have analysis and data collection carried out at the same time? Data analysis It is important to define the analysts’ profiles (physiotherapists?) and their experience with qualitative methodology. Once the framework was harmonized, did new codes emerge from data? has there been a reflective/iterative process? It is necessary to reflect on the positioning of the analysts and the resolution of possible biases: How have possible biases derived from the analyst been avoided? peer-review of conflicting texts? Patient or Public Involvement It should be relevant to mention that patients and members of the public were not involved in the validation of the data (validation techniques were not used after the analysis?). Results Were differences found between characteristics of participants and non-participants in the qualitative questions? Are there profiles that may have been lost? It would be clarifying to specify that there are two different coding

	frameworks, one per question. Is there any relevant difference (frequency or intensity) due to any of the sociodemographic characteristics in the discourses (age, gender, education, etc.)? Discussion It would be important to reflect more on possible sampling limitations, profiles not captured and motivation of the participants. Table 2-3: dark colours make reading difficult.

VERSION 1 – AUTHOR RESPONSE

REVIEWER #1
COMMENTS TO THE AUTHORS

This is an interesting study, using innovative methods for data collection, it is well written and provides valuable insights in how to handle shoulder pain patients. Before being published, there are however some minor corrections that I believe should be made.
Below you find a detailed list of comments on your manuscript.
AUTHORS' RESPONSE

We thank the reviewer for their positive comments.

COMMENTS TO THE AUTHORS

Strengths and limitations of the study:
Line 51. You state that you used "high-quality methods", which may be true, but in my opinion these are "standard procedures". If you thoroughly followed these standardized procedures it is a strength, but you should not claim that they are high-quality methods.
AUTHORS' RESPONSE

We have removed the phrase 'high-quality methods' from this dot point.

(Page 3, 2nd dot point)

- The online experiment which provided data for this study used randomisation and allocation concealment

COMMENTS TO THE AUTHORS

Introduction
Lines 97-101 in your Introduction is more of a discussion. I suggest that you remove this part.
AUTHORS' RESPONSE

We have removed this part as suggested. The paragraph now reads:

(Page 5, 2nd paragraph)

As part of our online randomised controlled experiment [14], we collected qualitative data that could help to uncover why preferences differed based upon the diagnostic label people received. The aim of this study was to explore how people with and without shoulder pain in our online experiment perceived different labels for rotator cuff disease in terms of words or feelings evoked by the label and treatments they feel are needed.

COMMENTS TO THE AUTHORS

Materials and methods
Participants and recruitment
Line 112. You mention Qualtrics, with no further explanation here (you do explain it later in text). I suggest that you provide this information to the reader here, when first mentioning Qualtrics.
AUTHORS' RESPONSE

We have moved the description of Qualtrics earlier in the below paragraph as suggested.

(Page 6, 2nd paragraph)

Participants aged 18-65 years old from Australia, New Zealand, United States, United Kingdom, and Canada were recruited through Qualtrics (www.qualtrics.com) between April and June 2020. Qualtrics is a market research company that recruits participants using existing, nationally representative panels of individuals who have previously agreed to complete surveys. Qualtrics employs random sampling and provides incentives for participants to complete surveys (e.g. cash, airline miles, gift cards). Details on the sampling and recruitment procedures Qualtrics use are reported elsewhere [14, 23]. Qualtrics recruited three groups of participants (evenly distributed) for our study: those who had never experienced shoulder pain, those who had shoulder pain at the time of participation, and those who had previously experienced shoulder pain but were pain-free at the time of participation.

COMMENTS TO THE AUTHORS

Data collection
Line 129. Participants were randomised to the six groups. You earlier described three groups of participants based on experience of shoulder pain – was this randomisation performed within each of these groups? From your results I guess that it was, but I suggest that you make this clear also in the Methods section. In line with this, figure 1 indicates that randomisation was not based on the three groups, which I find confusing. Please, clarify.
AUTHORS' RESPONSE

Randomisation was not based on the three groups of participants with differences experiences of shoulder pain. We have made this clear in the revised manuscript.

(Page 7, 2nd paragraph)

Participants read a vignette describing a patient with rotator cuff disease and were randomised to one of six labels. Randomisation was not stratified by the three groups of participants with different experiences of shoulder pain.

COMMENTS TO THE AUTHORS

Data analysis
Line 159. I have difficulties with your reference 24. I expected literature on Content analysis, for example Content analysis: an introduction to its methodology by Klaus Krippendorff. I suggest that you add this or a similar reference to the method applied.
AUTHORS' RESPONSE

Thank you for pointing out this error. We have fixed reference 24 now – it should have been: *Weber RP. Content analysis. 2nd ed. Thousand Oaks, California: Sage, 1990:117–24.*

COMMENTS TO THE AUTHORS

Line 160. I suggest that you provide some information about the coders here, that they both are experienced physiotherapists… (appears later, but inform the reader at this stage).
AUTHORS' RESPONSE

We have moved information about the experience of the coders earlier as suggested.

(Page 10, 2nd paragraph)

> Two researchers with experience in qualitative research and a physiotherapy background (JZ and ZAM) initially read through the responses to become familiar with their content. As such, the analysis represents the perspectives of physiotherapists currently working in research and with extensive experience managing patients with musculoskeletal pain. To develop the coding frameworks (one for each question), an inductive approach embedded in grounded theory was used. The two researchers independently coded 50 responses from each labelling group for both questions (~24% of all responses). The frameworks were then compared, discussed and harmonised into the one framework for each question for the next stage of coding.

COMMENTS TO THE AUTHORS

Results
I find the results clear and presented in a sufficient way.
AUTHORS' RESPONSE

We thank the reviewer for their positive comment.

COMMENTS TO THE AUTHORS

Discussion
Summary of key findings
Line 238. Your summary of key findings is relevant, but I find I too long, too much repetition of results. Try to shorten it, and do not repeat numbers and proportions that can be found in the Results section.
AUTHORS' RESPONSE

We have removed numeric values as suggested. However, we would like to keep the rest of the section as it is. We believe all information in this section is necessary to explain the findings of this

content analysis in context of the main quantitative findings from our randomised controlled experiment.

COMMENTS TO THE AUTHORS

Strengths and weaknesses
Line 265. You present some weaknesses, but did you consider the fact that your population is "selected". Only those registering in Qualtrics are eligible. Who are you missing? Older people? Poor people? I suggest that you reflect also on this limitation.
AUTHORS' RESPONSE

Although Qualtrics only recruits participants from panels of individuals who have previously agreed to complete online surveys ('market research panels'), we believe our sample is representative of people with shoulder pain in primary care.

The mean age of our sample (40.3 years old) is similar to other studies conducted in primary care (e.g. 42 years old from https://bmcmusculoskeletdisord.biomedcentral.com/articles/10.1186/1471-2474-12-119/tables/2). Our sample of participants with current or previous shoulder pain have similar healthcare utilisation compared to data from 2,497 patient encounters with a general practitioner in Australia from 2011-2016 (https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0227688&type=printable). For example:

- Imaging (our study vs. Australian sample): 44% vs. 43%
- Surgery (our study vs. Australian sample): 9% vs. 5% (data from the Australia sample was on 'referral to a surgeon')
- Injections (our study vs. Australian sample): 21% vs. 20%

Our sample of participants currently experiencing shoulder pain also has similar SPADI scores (Shoulder Pain and Disability Index; ~50/100) compared to trials included in Cochrane reviews of people with rotator cuff disease (e.g. https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD012224/full; https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD012225/full) and other studies in primary care (e.g. https://academic.oup.com/rheumatology/article/44/6/800/2899328).

We have now added this information to the revised manuscript.

(Page 14, 3rd paragraph)

> We were missing data from 318 participants who were randomised but did not complete outcome measures. However, our sample appears representative of people presenting with shoulder pain in primary care in terms of demographics, healthcare utilisation, and shoulder pain and function [3, 30-33].

COMMENTS TO THE AUTHORS

Comparison to literature
Lines 322-327. This is an interesting discussion, but you do not need all these numbers and proportions to understand the message. All these data can be found in the Results section, and the readability increases if you remove them from here.

AUTHORS' RESPONSE

We have removed these numbers as suggested.

COMMENTS TO THE AUTHORS

Conclusion
I agree with your conclusion, but I find it too long, and again repeating the results. Try to make it more distinct.

AUTHORS' RESPONSE

We have shortened the conclusion as suggested.

(Page 18, 2nd paragraph)

> Words or feelings evoked by certain labels for rotator cuff disease and perceived treatment needs may explain why some labels drive management preferences towards surgery and imaging more than others. Feelings of psychological distress and that the condition is serious and has a poor prognosis, and the need for treatment/investigation and surgery, were common among those labelled with a *rotator cuff tear* and *subacromial impingement syndrome,* but not among those labelled with *bursitis*. Interventions addressing misconceptions and perceived need for unnecessary care in patients given different labels for rotator cuff disease, and the clinicians who provide these labels, should be tested.

REVIEWER #2

COMMENTS TO THE AUTHORS

Study design
Has been used any qualitative methodology approach (phenomenology, hermeneutics, grounded theory)?

AUTHORS' RESPONSE

We have added the qualitative methodological approach to our methods.

(Page 10, 2nd paragraph)

> To develop the coding framework, an inductive approach embedded in grounded theory was used.

COMMENTS TO THE AUTHORS

It should be necessary to define the level of interpretation (descriptive, descriptive-interpretative, etc) and the perspective / level of participation (ETIC).

AUTHORS' RESPONSE

We have added more details about the interpretation and perspective of our content analysis.

(Page 10, 2nd paragraph)

> To develop the coding framework, an inductive approach embedded in grounded theory was used.

(Page 10, 2nd paragraph)

> Two researchers with experience in qualitative research and a physiotherapy background (JZ and ZAM) initially read through the responses to become familiar with their content. As such, the analysis represents the perspectives of physiotherapists currently working in research and with extensive experience managing patients with musculoskeletal pain.

COMMENTS TO THE AUTHORS

Has been considered sufficiency and pertinence of the sample based on the analysed qualitative data?

AUTHORS' RESPONSE

Our analysis was based on free-text data collected in a randomised controlled experiment of 1,308 participants. We have free-text data for all 1,308 participants and <3% of participants provided irrelevant responses (see Table 4). Hence, we have a sufficient amount of data for this content analysis.

COMMENTS TO THE AUTHORS

Have analysis and data collection carried out at the same time?

AUTHORS' RESPONSE

Data was collected between April and June 2020. Data was analysed between July and August 2020. We have added this to the revised manuscript.

(Page 5, 4th paragraph)

> Participants aged 18-65 years old from Australia, New Zealand, United States, United Kingdom, and Canada were recruited through Qualtrics (www.qualtrics.com) between April and June 2020.

(Page 10, 3rd paragraph)

The development and use of the frameworks occurred between July and August 2020.

COMMENTS TO THE AUTHORS

Data analysis
It is important to define the analysts' profiles (physiotherapists?) and their experience with qualitative methodology.

AUTHORS' RESPONSE

We have added this detail to the revised manuscript.

(Page 10, 2nd paragraph)

> Two researchers with experience in qualitative research and a physiotherapy background (JZ and ZAM) initially read through the responses to become familiar with their content. As such, the analysis represents the perspectives of physiotherapists currently working in research and with extensive experience managing patients with musculoskeletal pain. To develop the coding embedded in grounded theory, an inductive approach was used. The two researchers independently coded 50 responses from each labelling group for both questions (~24% of all responses). The frameworks were then compared, discussed and harmonised into the one framework for each question for the next stage of coding.

COMMENTS TO THE AUTHORS

Once the framework was harmonized, did new codes emerge from data? has there been a reflective/iterative process?

AUTHORS' RESPONSE

No new codes emerged from the data once the framework was harmonized. This was likely because the two reseachers initially read through the responses to become familiar with their content, and then independently coded a large proportion of responses (~24%) to compare and discuss before harmonising the framework. Once the framework had been harmonised/developed, the two researchers independently applied the framework to a random sample of responses, ensuring at least 20% of responses from each labelling group were coded. No further adjustments/iterations were needed since level of agreement between the two researchers coding a random sample of responses was 'almost perfect' for question 1 (range across the six labelling groups: k=0.90 to 0.97) and question 2 (k=0.91 to 0.97).

COMMENTS TO THE AUTHORS

It is necessary to reflect on the positioning of the analysts and the resolution of possible biases: How have possible biases derived from the analyst been avoided? peer-review of conflicting texts?

AUTHORS' RESPONSE

We have reflected on the positioning of the analysis in the methods.

(Page 10, 2nd paragraph)

> Two researchers with experience in qualitative research and a physiotherapy background (JZ and ZAM) initially read through the responses to become familiar with their content. As such, the analysis represents the perspectives of physiotherapists currently working in research and with extensive experience managing patients with musculoskeletal pain.

However, we acknowledge that professional bias and beliefs may have played a role in the development and application of the coding frameworks. We have now acknowledged this in the limitations.

(Page 15, 1st paragraph)

> Finally, since two researchers, both with a physiotherapy background developed and applied the coding framework, it is possible professional bias and beliefs may have influenced the coding.

COMMENTS TO THE AUTHORS

Patient or Public Involvement
It should be relevant to mention that patients and members of the public were not involved in the validation of the data (validation techniques were not used after the analysis?).
AUTHORS' RESPONSE

We have added this as suggested.

(Page 11, 2nd paragraph)

> Patients and members of the public were not involved in the design of this study nor were they involved in the validation of the data.

COMMENTS TO THE AUTHORS

Results
Were differences found between characteristics of participants and non-participants in the qualitative questions? Are there profiles that may have been lost?
AUTHORS' RESPONSE

We did not have data on the 318 participants that were randomised but did not complete outcome measures. However, we can demonstrate that our sample is representative of people with shoulder pain in primary care by comparing demographics, healthcare utilisation and disability to other samples.

The mean age of our sample (40.3 years old) is similar to other studies conducted in primary care (e.g. 42 years old from https://bmcmusculoskeletdisord.biomedcentral.com/articles/10.1186/1471-2474-12-119/tables/2). Our sample of participants with current or previous shoulder pain have similar healthcare utilisation compared to data from 2,497 patient encounters with a general practitioner in Australia from 2011-2016

(https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0227688&type=printable). For example:

- Imaging (our study vs. Australian sample): 44% vs. 43%
- Surgery (our study vs. Australian sample): 9% vs. 5% (data from the Australia sample was on 'referral to a surgeon')
- Injections (our study vs. Australian sample): 21% vs. 20%

Our sample of participants currently experiencing shoulder pain also has similar SPADI scores (Shoulder Pain and Disability Index; ~50/100) compared to trials included in Cochrane reviews of people with rotator cuff disease (e.g. https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD012224/full; https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD012225/full) and other studies in primary care (e.g. https://academic.oup.com/rheumatology/article/44/6/800/2899328).

We have now added this information to the revised manuscript.

(Page 14, 3rd paragraph)

> We were missing data from 318 participants who were randomised but did not complete outcome measures. However, our sample appears representative of people presenting with shoulder pain in primary care in terms of demographics, healthcare utilisation, and shoulder pain and function [3, 30-33].

COMMENTS TO THE AUTHORS

It would be clarifying to specify that there are two different coding frameworks, one per question.
AUTHORS' RESPONSE

We have added this information as suggested and pluralised 'framework' throughout the manuscript to make it clear there were more than one framework.

(Page 10, 2nd paragraph)

> To develop the coding frameworks (one for each question), an inductive approach embedded in grounded theory was used. The two researchers independently coded 50 responses from each labelling group for both questions (~24% of all responses). The frameworks were then compared, discussed and harmonised into one framework for each question for the next stage of coding.

COMMENTS TO THE AUTHORS

Is there any relevant difference (frequency or intensity) due to any of the sociodemographic characteristics in the discourses (age, gender, education, etc.)?
AUTHORS' RESPONSE

We did not conduct any sub-group analyses as part of this content analysis because no sub-group analyses showed any difference for our quantitative outcomes (https://www.jospt.org/doi/full/10.2519/jospt.2021.10375). Further, since participants' free-text responses could contain up to nine codes, people with certain characteristics that expressed more than one code/theme in their response would be over-represented in any sub-group analysis we perform.

COMMENTS TO THE AUTHORS

Discussion
It would be important to reflect more on possible sampling limitations, profiles not captured and motivation of the participants.

AUTHORS' RESPONSE

Although Qualtrics only recruits participants from panels of individuals who have previously agreed to complete online surveys ('market research panels'), we believe our sample is representative of people with shoulder pain in primary care (see response to earlier comment).

We did not collect data on the motivation of participants so cannot comment on this.

COMMENTS TO THE AUTHORS

Table 2-3: dark colours make reading difficult.

AUTHORS' RESPONSE

We have changed the colour scheme to avoid dark colours.